Cell, Vol. 62, 945–954, September 7, 1990, Copyright © 1990 by Cell Press

Role of Eukaryotic-Type Functional Domains Found in the Prokaryotic Enhancer Receptor Factor σ^{54}

Selina Sasse-Dwight and Jay D. Gralla
Department of Chemistry and Biochemistry and the Molecular Biology Institute
University of California, Los Angeles
Los Angeles, California 90024

Summary

E. coli σ^{54} protein confers on promoters containing its recognition sequence the ability to be activated from distant DNA sites. Its functional domains include two leucine zipper motifs, an acidic region, and a glutamine-rich domain. Several domains were disrupted and the assembly of mutant transcription complexes was probed in vivo by footprinting. Promoter recognition was seen to depend on a C-terminal region containing a prokaryotic helix-turn-helix motif. Within the resulting stable closed complex, two leucine zipper motifs assist in positioning the σ^{54} polymerase near the DNA region that must be melted upon activation. Finally, DNA opening depends on the σ^{54} acid domain. The uncoupling of promoter recognition from DNA melting, mediated by the unusual domain structure of this prokaryotic protein, may be responsible for σ^{54}'s ability to mediate activation from distant sites.

Introduction

The *glnA*p2 promoter was the first prokaryotic promoter shown to contain upstream regulatory elements that could be moved more than a kilobase away and still mediate transcriptional activation (Reitzer and Magasanik, 1986). The upstream enhancer-like elements are the binding sites for the NR_I activator protein (NtrC), which is activated by phosphorylation in response to low nitrogen levels (Ninfa and Magasanik, 1986; Keener and Kustu, 1988). Phosphorylated NR_I normally acts from ~100 bp upstream of the transcription start site to activate a prebound but inactive RNA polymerase. This activation causes the polymerase to melt the promoter DNA at the transcriptional start site both in vitro and in vivo in what has been determined to be the slow step in transcriptional activation (Sasse-Dwight and Gralla, 1988; Popham et al., 1989). The polymerase holoenzyme contains σ^{54}, a relatively minor σ factor, rather than the normal σ^{70} factor. This $E\sigma^{54}$ holoenzyme (the complex containing σ^{54} and core RNA polymerase) serves as the receptor protein complex for receipt of the melting signal from the remotely bound activator protein.

Although σ^{54} is a rare protein, it is the required receptor protein in the only known examples where activator elements can be moved far from bacterial promoters in Escherichia coli and retain their function (Reitzer and Magasanik, 1986; Birkmann and Böck, 1989; Gralla, 1989a). The activators involved in those E. coli σ^{54}-dependent promoters that have been studied, *glnA*p2 and *fdh*, are genetically and physiologically distinct (Birkmann et al., 1987). Thus, it appears to be the common involvement of σ^{54} that allows these promoters to be activated from different distant sites by different proteins. σ^{54} thus appears to mediate a eukaryotic-type phenomenon in a prokaryotic organism.

Below, the sequence of the E. coli σ^{54} gene is determined and compared with its analogs from three different nitrogen-fixing bacteria (Merrick and Gibbins, 1985; Merrick et al., 1987; Ronson et al., 1987). This reveals conserved motifs of the type more commonly associated with eukaryotic transcription factors. These include two leucine zipper motifs (Landshulz et al., 1988; Vinson et al., 1989), a strongly acidic region (Giniger and Ptashne, 1987; Hope et al., 1988; Brendel and Karlin, 1989), and a glutamine-rich region (see Courey and Tjian, 1988). Pairs of eukaryotic leucine zippers are required for transcription and have been shown to cooperate to form a DNA binding domain through a coiled-coil helical structure (O'Shea et al., 1989a, 1989b; Sorger and Nelson, 1989). Glutamine-rich and acidic regions function as activation domains in certain eukaryotic factors. The mechanism by which they activate transcription is not known.

Thus, σ^{54} mediates a eukaryotic-type phenomenon in E. coli and contains domains more characteristic of eukaryotic than prokaryotic transcription factors. The protein actually appears to resemble a hybrid factor since it also contains potential helix-turn-helix (HTH) domains (Merrick and Gibbins, 1985; Merrick et al., 1987) that are well known to mediate DNA binding in bacteria and, in modified forms, in higher cells. Below we take advantage of bacterial genetics to introduce deletions into several of these domains and study the altered proteins in cells lacking the wild-type σ^{54}. In vivo footprinting is used to study the resulting defective transcription complexes along the chromosomal *glnA*p2 promoter. The results lead to proposals for the functions of several of the critical domains in enhancer-dependent activation in E. coli.

Results

σ^{54} Contains Domains Similar to Eukaryotic Transcription Factors

The E. coli σ^{54} gene (*rpoN*, *glnF*) from plasmid pTH7 (Hunt and Magasanik, 1985) was sequenced (see Figure 8), and the resulting amino acid sequence was compared with the previously determined σ^{54} amino acid sequences from other bacteria (Merrick and Gibbins, 1985; Merrick et al., 1987; Ronson et al., 1987). The sequences are on average 25% identical (42% conserved). The areas of strongest sequence conservation are those previously termed region I and region III (Merrick et al., 1987) (see Figure 1). The σ^{54} proteins are not similar to the main family of bacterial σ factors, however, which includes the major E. coli σ factor σ^{70} (Merrick et al., 1987).

Interestingly, several of the conserved regions are similar in sequence composition and periodicity to domains known to be important for eukaryotic transcription factors.

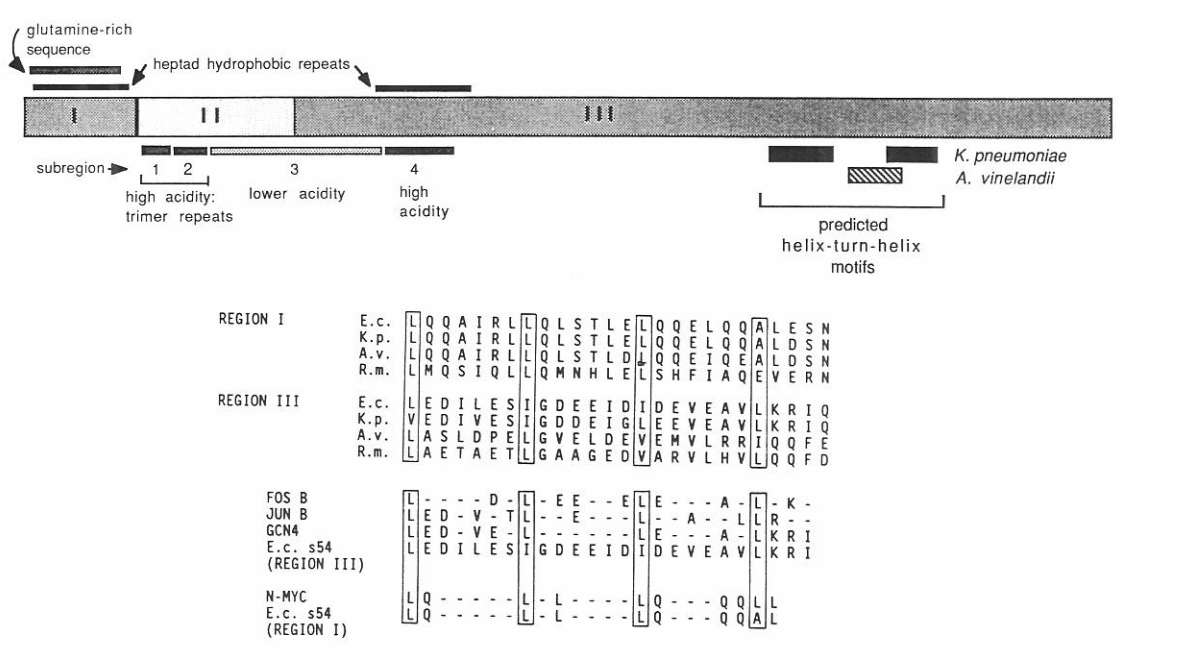

Figure 1. Schematic Diagram of the σ^{54} Protein Illustrating the General Regions of Conservation (I–III) and Its Conserved Motifs

The actual sequences of the two conserved heptad hydrophobic repeats found in the E. coli σ^{54} protein (marked above regions I and III) and their comparison to the homologous σ^{54} sequences from three different bacteria and to eukaryotic proteins (reviewed by Busch and Sassone-Corsi, 1990) are shown below the schematic diagram. In these comparisons, the residues involved in the hydrophobic heptad repeat are boxed. In addition to the hydrophobic heptad repeats, the glutamine-rich sequences and the two potential HTH motifs predicted based on other σ^{54} sequences are indicated by bars on the schematic diagram. The acidic region, composed of four subregions, is also illustrated. The dark stippling of regions I and III compared with region II indicate higher overall conservation of amino acid sequence.

Figure 1 shows the sequences of two potential leucine zipper repeats in σ^{54} in comparison to other members of the σ^{54} family and to eukaryotic proteins. The region III hydrophobic heptad is conserved among the bacterial species and most resembles yeast GCN4 (36% identity). In addition, it contains part of a highly acidic domain that begins within region II. Although this domain is poorly conserved in sequence, it is much more conserved with respect to acidity, giving the entire acidic domain of the E. coli protein a total net charge of −37. In contrast to region III, the conserved region I repeat most resembles the glutamine-rich leucine zipper of N-Myc (39% identity). The glutamines make up 25% of the residues in this region as was also found for the glutamine-rich sequences in human Sp1 (Courey and Tjian, 1988; Courey et al., 1989). The C-terminus of σ^{54} is also moderately conserved and contains various HTH motifs (see Figure 1) (Merrick et al., 1987).

The initial strategy in characterizing these various domains was to take advantage of genetic screens to isolate mutants that express altered forms of σ^{54} in cells lacking the wild-type protein. The E. coli σ^{54} gene was carried on the plasmid pTH7, a derivative of pBR322, under the control of the *tac* promoter (Hunt and Magasanik, 1985). Mutations were constructed either by using nuclease Bal31 to delete from a unique restriction site within a domain or via oligonucleotide-directed deletion mutagenesis. The pTH7 mutant DNA library containing small deletions was then transformed into a strain unable to produce σ^{54} from its chromosome. The resulting colonies were subjected to two genetic screens to identify expressed mutants that were nonfunctional with respect to their ability to support expression of the *glnALG* promoter.

The first of the two screens ensures that the mutant protein is expressed from the vector and is functional with respect to at least one criterion. When the wild-type σ^{54} protein is overexpressed due to induction with IPTG, growth of the cells is severely inhibited, resulting in tiny colonies. This was observed previously upon overexpression of the Bacillus subtilus SpoIIAC σ factor and was suggested to be due to competition of this protein with other σ factors for core RNA polymerase (Yudkin, 1987). Whatever the cause of toxicity upon overexpression, it serves as a convenient phenotype for assaying the functional stability of the σ^{54} protein; as with the wild-type protein, only those mutant proteins that are actually expressed within the cell are able to cause toxicity upon IPTG induction.

The colonies containing proteins that were expressed by the above criterion were then screened for their ability to activate the *glnA*p2 promoter. This involved plating (without high IPTG) the mutants on W-Arg medium, a minimal medium on which *glnALG* expression and therefore functionally wild-type σ^{54} is required for growth. Nongrowing strains that carried defective forms of σ^{54} were picked from replica plates on rich media, and their plasmids were isolated and sequenced within the region targeted for mutagenesis.

Following DNA sequencing of the mutant plasmids the

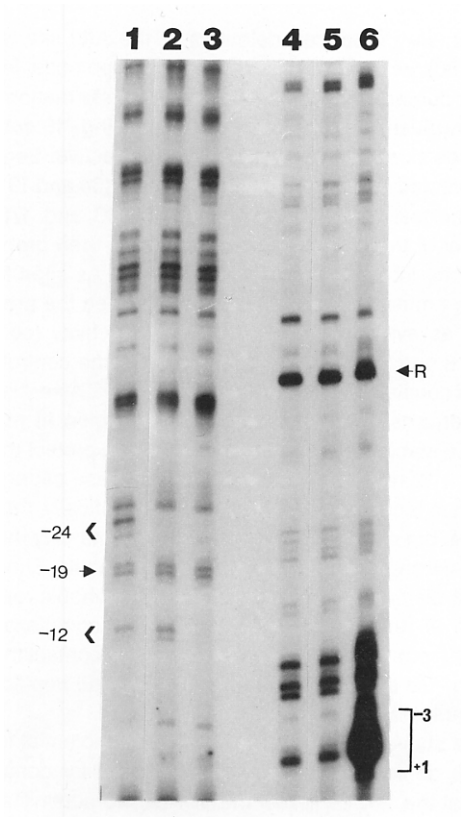

Figure 2. Probing Region I Mutant Transcription Complexes In Vivo

The autoradiograph shows DMS (lanes 1–3) and $KMnO_4$ (lanes 4–6) patterns at the *glnAp2* promoter region in vivo. Lanes 1 and 4, control in which no plasmid DNA, and therefore no σ^{54} protein, is present in the cell; lanes 2 and 5, the mutant plasmid Sal58 (residues 18–31 deleted) is present in the cell as the only source of σ^{54}; lanes 3 and 6, the wild-type pTH7 plasmid DNA present in the cell. The −12 and −24 regions and the −19 band are marked with arrows for the DMS samples (lanes 1–3) and the −3 to +1 region is marked with a bracket for the $KMnO_4$ samples (lanes 4–6).

functional defect associated with the mutant σ^{54} proteins was determined in vivo by use of primer extension footprinting (Gralla, 1985). There are two complexes that can be assayed by this method. Recall that under nonactivating conditions σ^{54} associates with core RNA polymerase and directs assembly of an inactive, closed complex covering the *glnA* transcription startsite. This complex can be detected by protection from dimethylsulfate (DMS) attack of critical guanines within the −12 and −24 regions of the promoter (Sasse-Dwight and Gralla, 1988; Popham et al., 1989). Additionally, the ability of the polymerase to respond to activation by melting the DNA can be assayed by testing for potassium permanganate ($KMnO_4$) sensitivity of the region around the startpoint (Sasse-Dwight and Gralla, 1988, 1989). Both of these in vivo assays were previously established using wild-type σ^{54} (Sasse-Dwight and Gralla, 1988), and the results obtained agreed well with in vitro conclusions (Ninfa et al., 1987; Popham et al., 1989). The procedure used in these experiments requires a slight modification of the original extension protocol, since the promoter is probed on chromosomal rather than plasmid DNA. We therefore used a modification of the polymerase chain reaction (PCR) to amplify and clarify the signal (see Experimental Procedures; Axelrod and Majors, 1989).

The Heptad Hydrophobic Repeats Are Both Required for Activation

Nuclease Bal31 was used to delete from the unique SalI restriction site (amino acid 28), which interrupts the leucine-rich hydrophobic heptad repeat of region 1. Following transformation of the mutant plasmids into a host strain lacking wild-type σ^{54}, only 2 colonies out of 64 tested were found to express phenotypically stable σ^{54} proteins that were unable to induce the *glnALG* operon in the cell. These mutants (Sal57 and Sal58) carried in-frame deletions of amino acids 20–31 and 18–31, which remove over one-half of the region involved in the leucine heptad repeat.

These two mutants were assayed in vivo for their ability to open the promoter DNA and expose the transcription start point. Lanes 4–6 of Figure 2 show the pattern obtained when potassium permanganate is used to modify the chromosomal promoter DNA in vivo. The drug rifampicin was added to the cells for 5 min prior to addition of $KMnO_4$ to trap any open complexes that might form (Sasse-Dwight and Gralla, 1988). When wild-type σ^{54} is expressed, the bands from −3 to +1 are strongly hyperreactive to attack, as evidenced by the strong permanganate sensitivity of the chromosomal signal in lane 6. In contrast, when the Sal58 region I deletion mutant is instead present in the cell, the low reactivity at the transcriptional start site is almost identical (lane 5) to that seen when σ^{54} is completely absent (lane 4). That is, no DNA open promoter complexes are observed. An identical pattern was observed for the Sal57 mutant (data not shown). Thus, the mutant Eσ^{54} transcription complexes are not able to melt the promoter DNA under activating conditions in vivo.

This inability to form an open complex in vivo could be due to any number of defects, including an inability of the mutant protein to associate with the promoter in a closed complex. This can be tested by probing for protection of the −12 and −24 promoter regions with DMS. It was shown previously that protection of these two critical promoter regions requires σ^{54} and occurs in both closed and open complexes (Sasse-Dwight and Gralla, 1988).

The results of using DMS to footprint chromosomal *glnAp2* mutant transcription complexes are also shown in Figure 2. As controls on the PCR-modified technique, lanes 1 and 3 show the DMS patterns obtained when the promoter is probed in the absence of σ^{54} (no plasmid DNA present) or the presence of wild-type σ^{54}, respectively. As observed previously on plasmid DNA (Sasse-Dwight and Gralla, 1988), when σ^{54} is absent the bands at both the −12 and −24 regions (marked by arrows) are approximately equal in intensity to one another and to the −19 band (lane 1). However, when σ^{54} is present these two bands are both much lower in intensity than the −19 band due to protection of the −12 and −24 transcription elements (lane 3).

The in vivo footprinting results obtained for the Sal58 mutant using DMS (lane 2) shows that the mutant σ^{54}

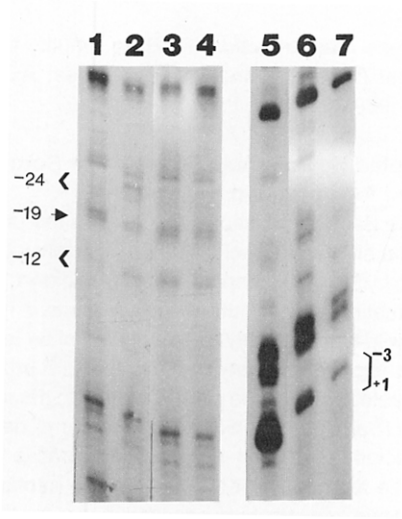

Figure 3. Probing Region III Mutant Complexes In Vivo

The autoradiograph shows DMS (lanes 1–4) and $KMnO_4$ (lanes 5–7) patterns obtained in vivo at the *glnAp2* downstream promoter region in the absence of any σ^{54} in the cell (lane 2) and the presence of either wild-type (lanes 1 and 5) or region III deletion mutant σ^{54} proteins (lanes 3, 4, 6, 7) in the cell. Lanes 3 and 6, in vivo patterns obtained when the mutant plasmid Afl2 (residues 174–190 deleted) is present in the cell. Lanes 4 and 7, in vivo patterns obtained when the mutant plasmid Afl7 (residues 178–183 deleted) is present in the cell. The −12 and −24 regions as well as the −19 band are marked by arrows for the DMS samples, and the −3 to +1 region is bracketed for the $KMnO_4$ samples.

polymerase does indeed occupy the promoter DNA inside the cell. This is evidenced by the strong protection of the band marking the promoter −24 element. Specifically, this band is much lower in intensity than the reference band at −19, just as was observed for the wild-type protein. Repeated experiments indicated that although the mutant holoenzymes cannot melt and activate the DNA, they bind tightly to it in vivo, as evidenced by −24 protection.

The mutant transcription complexes, however, do differ from wild-type in one interesting respect. The experiment shows that the mutant Sal58 fails to protect the promoter −12 region band that lies just adjacent to the region to be melted (lane 2). Unlike in the wild-type control (lane 3), this band is equal in intensity to the band at −19, as observed when σ^{54} was absent from the cell (lane 1). Because protection was seen at −24, the overall pattern for Sal58 (lane 2) and Sal57 (data not shown) is unlike the pattern obtained either in the presence (lane 3) or absence (lane 1) of wild-type protein and indicates that the mutant protein binds only to −24 and not to −12.

The above results demonstrate that the leucine repeat domain of region I is not required for σ^{54} to bind to the promoter DNA in vivo, as evidenced by complete protection of the −24 region, but is required for contacts to form at −12. Thus, the mutant $E\sigma^{54}$ complex is able to form a complex at the promoter, but this is defective, as evidenced by an inability to make contacts near the transcriptional start site. The defective yet bound transcription complex cannot melt and activate the promoter.

Next, we introduced deletions at the AflII site (amino acid 180), which falls within the second potential leucine zipper domain in region III. Following transformation of the mutant plasmid DNA library and screening, 16 out of 47 colonies screened as expressed but defective. Sequencing revealed deletions between residues 164 and 191, with the shortest removing residues 175–180 and 178–183. Several of these region III mutants were then probed as above for interactions with the promoter. As seen for the region I mutants, these mutants fail to melt the promoter DNA, as evidenced by low $KMnO_4$ reactivity (compare lanes 6 and 7 of Figure 3 with the wild-type control, lane 5) and consistent with their lack of function. Also, like their counterparts in region I, each of the region III mutants forms a stable closed complex that fails to protect the promoter −12 region in vivo (examples of these patterns are shown in lanes 3 and 4 of Figure 3). Specifically, the band at −24, but not at −12, is much lower in intensity than the −19 reference band for each of these mutants. Thus, mutations eliminating a portion of the hydrophobic repeat in region III result in mutant transcription complexes that bind the promoter DNA at −24 but do not contact the −12 region, just as was observed for the region I heptad deletion mutants.

The above data show that both of the potential leucine zipper domains are specifically required for contacts to form at the −12 region of the *glnAp2* promoter. This suggests that the two domains cooperate to induce formation of the structure that contacts the −12 region DNA. Based on studies showing that two leucine zipper motifs can cooperate to form a DNA binding domain by dimerizing into a coiled-coil structure (O'Shea et al., 1989b; Vinson et al., 1989), we infer that these two σ^{54} domains likely interact with one another to form such a structure. This intramolecular form of the leucine zipper protects the DNA element just adjacent to the region that must be melted for promoter activation to occur. The data show that when this structure is not positioned over the −12 region due to a mutation in σ^{54}, the promoter cannot be melted. Thus, the two hydrophobic heptad repeat domains appear to cooperate to form a structure that positions the activation domain.

The C-Terminal Region Helps Bring $E\sigma^{54}$ to the Promoter

These data show that the leucine zipper domains are not required to bring the σ^{54} polymerase to the DNA as evidenced by full protection of the promoter −24 region when either heptad repeat is mutated. A likely candidate for promoter recognition and binding is the HTH region (reviewed in Brennan and Matthews, 1989). Deletions were constructed to overlap the most C-terminal HTH motif (centered at position 393; Merrick and Gibbins, 1985) as well as a nearby alternative HTH motif. Following transformation with the mutant DNA library, several colonies were obtained that exhibited toxicity upon overexpression of σ^{54} and were specifically deficient in expression of the *glnALG* operon. Among these were two that extended only into the C-terminal HTH motif, indicating that this region is required for expression of the *glnALG* operon. In vivo

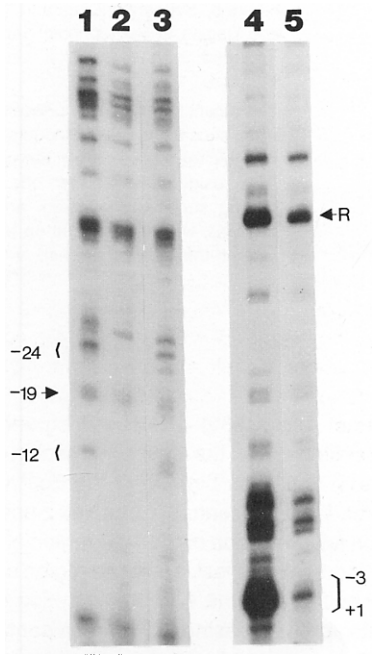

Figure 4. Probing Helix-Turn-Helix Mutant Complexes In Vivo

Lane 1, in vitro DMS pattern along naked DNA; lanes 2 and 4, in vivo DMS pattern obtained when wild-type σ^{54} (pTH7) is present in the cell; lanes 3 and 5, in vivo DMS patterns obtained when the HindIII mutant deletion plasmid A24 (residues 394–422 deleted) is present in the cell. The bands in regions −12 and −24 are marked with arrows as well as the band at −19, which serves as a reference for the DMS samples. In addition, the bands from −3 to +1 are marked for the $KMnO_4$ samples.

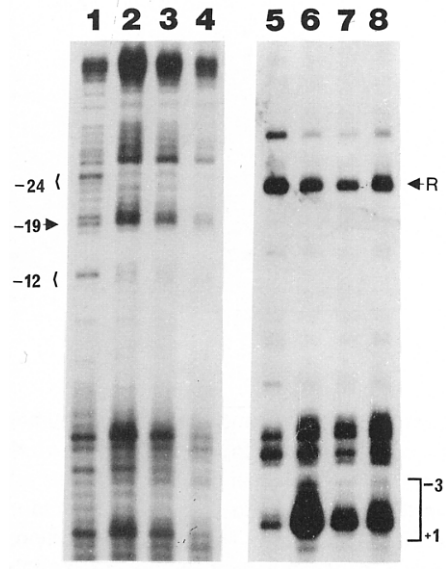

Figure 5. In Vivo DMS and $KMnO_4$ Patterns Obtained for the Acid Deletion Mutant

In vivo DMS, lanes 1–4; in vivo $KMnO_4$, lanes 5–8. Residues 51–77 were deleted from the acid deletion mutant. Lanes 2 and 6, in vivo patterns obtained when the wild-type σ^{54} protein is expressed in the cell. Lanes 1 and 5, in vivo patterns obtained when no σ^{54} is present in the cell. Lanes 3 and 7, in vivo patterns obtained when the acid deletion mutant Acdel9 is present in the cell. Lanes 4 and 8 show the patterns obtained for an identical mutant, Acdel10. The −12 and −24 regions are marked with an arrow for the DMS lanes. In addition, the bands from −3 to +1 and the reference band, R (see text), are marked next to the permanganate samples.

footprinting with $KMnO_4$ demonstrated that none of these mutants was able to melt the promoter as judged by lack of hypersensitivity at the transcriptional start site (see Figure 4, lane 5, for an example of the $KMnO_4$ pattern obtained for these mutants).

We next probed with DMS in vivo to learn whether the cause of this loss of activity was an inability of the σ^{54} polymerase to associate with the DNA. A typical example of such an HTH region deletion mutant pattern is illustrated in Figure 4, lane 3. The bands at both the −12 and −24 regions remain approximately equal in intensity to one another and to the −19 band for each of the mutants, resembling the in vitro pattern in which no σ^{54} is present (lane 1). Recall that each of the mutant proteins is functional with respect to its ability to induce toxicity and is therefore expressed in the cell. The result differs from either of the leucine zipper mutants where −24 region contacts were fully protected and mediated full promoter occupancy even without −12 binding. The additional loss of the −24 region contacts by the C-terminal deletions suggests that the HTH helps bring the polymerase to the DNA through contacts involving the −24 region.

Partial Deletion of the Acidic Residues Specifically Affects Melting

Region II contains a cluster of acidic residues arranged in four contiguous subregions conferring a net charge of −37 to the region as a whole (see Figure 1). Starting from the C-terminal side, the first subregion overlaps a portion of the leucine zipper in region III. (The small region III deletions did not significantly reduce the net charge of this subregion.) The next area represents a long subregion of lower average acidity followed by two subregions where the acidic residues are arranged in trimer repeats involving six residues each. As a result of its large size and overlap with critical leucine zipper residues, it was not desirable to delete this entire region. Instead, the trimer repeat subregions (amino acids 51–77) were deleted using oligonucleotide-directed mutagenesis, and two identical clones were tested for their phenotypes. These partial acid deletions still exhibited toxicity upon IPTG induction, indicating that they were expressed, and in contrast to the mutants described above, grew on arginine plates, indicating that σ^{54} function was at least residually intact. However, when these mutants were transformed into a reconstructed *glnA*–lacZ fusion strain deficient in σ^{54}, they produced a minimum of four times less β-galactosidase than wild-type transformants (unpublished data).

To determine whether the partial acid deletion mutant could melt the DNA to the same extent as wild-type, the mutants were probed with permanganate in vivo (Figure 5). The open complex signal for both of the separately isolated identical acid deletion mutants (lanes 7 and 8) is much lower than that observed for the wild-type σ^{54} (lane 6). To correct for variations in the $KMnO_4$ reactivity be-

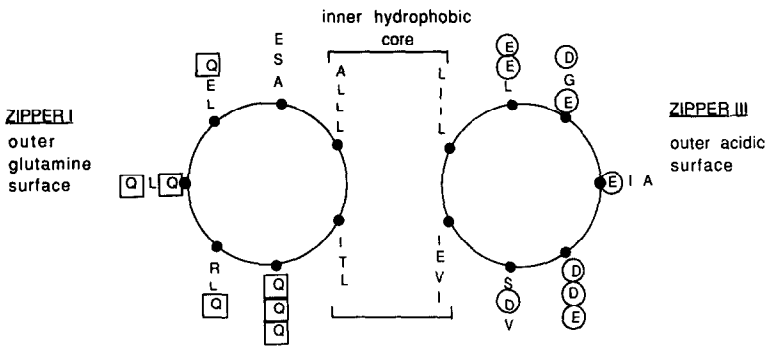

Figure 6. Schematic Diagram of the Region I and Region III Hydrophobic Heptad Repeats Arranged in a Coiled-Coil Leucine Zipper Structure

Two potential activating surfaces that appear when such a structure is modeled are marked on their respective zipper halves. The glutamine residues, which form one potential activating surface, are boxed, and the acidic residues, which appear as an activating surface on the other zipper half, are circled.

tween samples we compared the ratio of the intensity of the open complex signal to the intensity of a reference band, marked R, for each of the samples. This ratio is at least five times less for the acid deletion mutants compared with the wild-type. Thus, partial deletion of the acidic region results in a complex that can melt the DNA only partially.

As with the mutants in other regions, the inability to melt the DNA to the same extent as wild-type could be due to defective binding at the −12 or −24 region of the promoter, which can be tested with DMS probing. Lanes 3 and 4 of Figure 5 represent the results for the two identical acid deletion mutants. The pattern observed appears to be identical to the pattern obtained for the wild-type $E\sigma^{54}$ complex (lane 2); that is, the bands at −24 and −12 are much lower in intensity than the reference band at −19. This is not the case when no σ^{54} is present in the cell (lane 1), where the −24 and −12 bands are prominent in the pattern. Thus, the acid deletion mutants occupy the promoter DNA in the cell, demonstrating that the cause of the 5- to 10-fold loss of melting is not 5- or 10-fold less transcription complex formation. Rather, the acid region appears to act as a pure melting domain since its disruption interferes with DNA melting by a fully formed closed promoter complex.

Discussion

These experiments employed chemical probing of transcription complexes formed with mutant σ^{54} proteins in vivo to explore the roles of protein domains. These include two potential leucine zipper motifs, a highly acidic region, and an HTH motif. Each of these domains is shown to be required at a specific step for σ^{54} to induce DNA melting at the transcriptional start site in response to an activation signal from the NR_I (NtrC) enhancer.

Promoter Recognition and Binding

The E. coli σ^{54} protein must interact with two proteins (core RNA polymerase and NR_I phosphate) and the *glnA*p2 promoter DNA in the process of activation. Binding to core RNA polymerase occurs first, as in vitro data indicate that σ^{54} cannot bind promoter DNA in the absence of core (Ninfa et al., 1987). Once this holoenzyme forms, it must recognize and form a closed complex at the promoter (Ninfa et al., 1987; Sasse-Dwight and Gralla, 1988; Popham et al., 1989). The above experiments indicate that deletion within the C-terminal region containing an HTH motif abolishes binding of the $E\sigma^{54}$ complex to the promoter. This is in contrast to leucine zipper mutants, which retain full protection of the −24 region although the −12 region contacts are lost. Thus, the region carrying the HTH motif helps bring the $E\sigma^{54}$ transcription complex to the promoter, probably using −24 region contacts. Other regions could conceivably help with this, including a nearby moderately basic region.

Activation Involving Two Leucine Zippers and an Acidic Region

After $E\sigma^{54}$ forms a closed complex with the promoter it must be prepared to receive an activation signal from the phosphorylated enhancer protein NR_I bound upstream. Inactivation of either of the two separated leucine zippers does not prevent the closed complex from forming, as judged by full protection of the −24 region, but blocks it from contacting the −12 element and melting the DNA in response to the NR_I signal. Since the lost −12 contacts mark the beginning of the region that is normally melted, it seems that the two regions cooperate to form a structure that positions the melting domain.

By analogy to studies of other leucine zipper motifs (O'Shea et al., 1989a, 1989b; Sorger and Nelson, 1989; Vinson et al., 1989), the structure formed by cooperation of these two regions is presumed to be a coiled coil. Dimerization between various eukaryotic transcription factors such as *fos* and *jun* has been shown to occur through such a structure and results in the formation of a DNA binding domain (O'Shea et al., 1989b). In addition, 14 of the 15 residues that form the inner hydrophobic core of the predicted σ^{54} intramolecular leucine zipper are among the minority of amino acids that predominate in hydrophobic cores (Bowie et al., 1990). These regions also resemble dimerization motifs of the helix-loop-helix class of eukaryotic regulatory proteins, but σ^{54} does not dimerize either alone or in the presence of core RNA polymerase (Hunt and Magasanik, 1985). Thus, we propose that the two σ^{54} regions dimerize to form an intramolecular leucine zipper that is positioned just adjacent to the region of DNA to be melted. This positioning is required for the DNA melting that accompanies activation.

This intramolecular leucine zipper positioning domain has other features that are reminiscent of activation

regions in eukaryotic transcription factors. Each region is predicted to form a strongly amphipathic helix (Giniger and Ptashne, 1987; Hope et al., 1988), the polar surfaces of which present residues whose abundance marks many mammalian activation domains (see Figure 6). Specifically, the region I helix has a polar surface that contains 7 glutamines out of 12 residues (see Courey et al., 1989). The region III helix has an outer surface that contains 8 acidic residues out of 15 (see above references). Thus, the leucine zipper positioning domain presents two potential activating surfaces when associated with the closed complex that is awaiting activation. The roles of these two surfaces are not yet known. Since the residues are interdigitated with the critical hydrophobic residues and are numerous, extensive mutagenesis will be required to learn their function.

In strong contrast, deletion of region II acidic residues interferes with DNA melting without disturbing the ability of σ^{54} to form closed complexes that recognize both the −24 and −12 regions of the promoter. Thus, the acid residues can be considered to constitute a DNA melting domain. A possible model for how the melting occurs is discussed below.

Domain Organization in Comparision with σ^{70} and Eukaryotic Promoters

Both σ^{54} and the major σ factor σ^{70} (Gardella et al., 1989; Siegele et al., 1989) are similar in that they contain two separate regions that are required for recognition of two distinct promoter elements. For both proteins, one of the two regions has an HTH motif and helps binding to the DNA element farthest from the transcriptional start site, while the other does not have an HTH motif and contacts the closer element. However, the sequences recognized by these domains and the domains themselves are quite different (Reitzer and Magasanik, 1986; Gardella et al., 1989; Siegele et al., 1989). σ^{54} of course differs in that it has several functional domains resembling those of eukaryotic transcription factors. These domains probably explain why σ^{54}-dependent E. coli promoters can respond to remote activation as do eukaryotic promoters (Reitzer and Magasanik, 1986; Gralla, 1989a).

The E. coli *glnA*p2 transcription complex is very simple and can be reconstituted using only σ^{54}, phosphorylated enhancer NR_I, and the simple core polymerase. This is in obvious contrast to the eukaryotic complexes where many different factors participate in transcriptional activation (Nakajima et al., 1988; Buratowski et al., 1989). Nevertheless, the *glnA*p2 and the eukaryotic complexes have certain domains in common. These transcriptional domains are normally found on separate factors in eukaryotes but are delivered by a single protein, σ^{54}, in the bacterial complex. The difference in arrangement between these two systems may be due to the fact that the more complex eukaryotic machinery requires that many of these domains be donated by separate proteins to increase the diversity of transcriptional regulation.

Thus far, σ^{54} promoters are the only E. coli promoters known that allow activation from distant enhancer sites (Reitzer and Magasanik, 1986; Birkmann and Böck, 1989; Gralla, 1989a). *glnA*p2 also differs from typical E. coli promoters in that fully stable closed transcription complexes form. In contrast, most E. coli promoters are characterized by unstable closed complexes that occur only as transient intermediates (Gralla, 1989b). The preassembly of a stable but inactive partial transcription complex superficially resembles proximal assemblies at eukaryotic promotors (see Nakajima et al., 1988) and allows the *glnA* system to mediate separately the processes of promoter recognition and activation.

It may be that it is this unusual ability to uncouple stable promoter recognition from promoter activation that allows σ^{54} promoters, but not σ^{70} promoters, to be strongly activated from distant sites. Long-range assistance in promoter recognition of the type required by σ^{70} would require tight activator–RNA polymerase interactions, which could be difficult to break during transcription initiation and could tie up free RNA polymerase in undesirable complexes. This would be especially deleterious in a bacterium that must respond flexibly to changes in the external environment. For the σ^{54} eukaryotic-type mechanism, the distantly bound proteins need only touch the polymerase transiently in its prebound form to trigger it to melt the DNA.

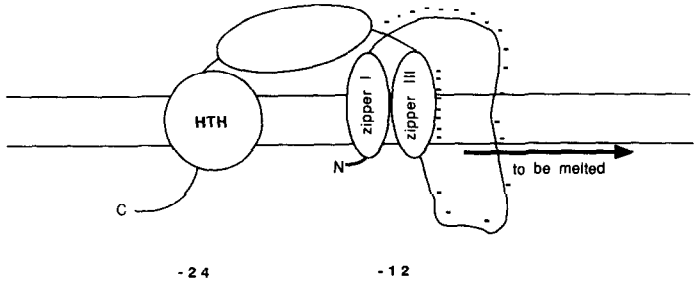

Figure 7. Cartoon of the σ^{54} Polymerase in a Closed Promoter Complex with the *glnA*p2 Promoter, Where the Core RNA Polymerase Is Invisible

This diagram illustrates the positioning of the HTH, which is essential for promoter recognition and closed complex formation, over the −24 region of the DNA and the cooperation of the two hydrophobic heptad repeats (labeled zipper I and III to correspond to the regions in which they are located) in forming a leucine zipper structure that contacts the −12 region. The positioning of surrounding regions of the protein is as predicted based on this parallel coiled-coil leucine zipper structure. For instance, the acidic domain, composed of the activating surface along the region III zipper half, the trimer repeats within region II, and the remaining acidic portion of region II, are predicted to lie within reach of the region of DNA to be melted upon open complex formation. The relative degree of negative charge within each subregion of the acidic domain is roughly proportional to the number of negative signs (-) shown. As illustrated in Figure 6, the glutamine-rich sequence is found on the region I zipper half.

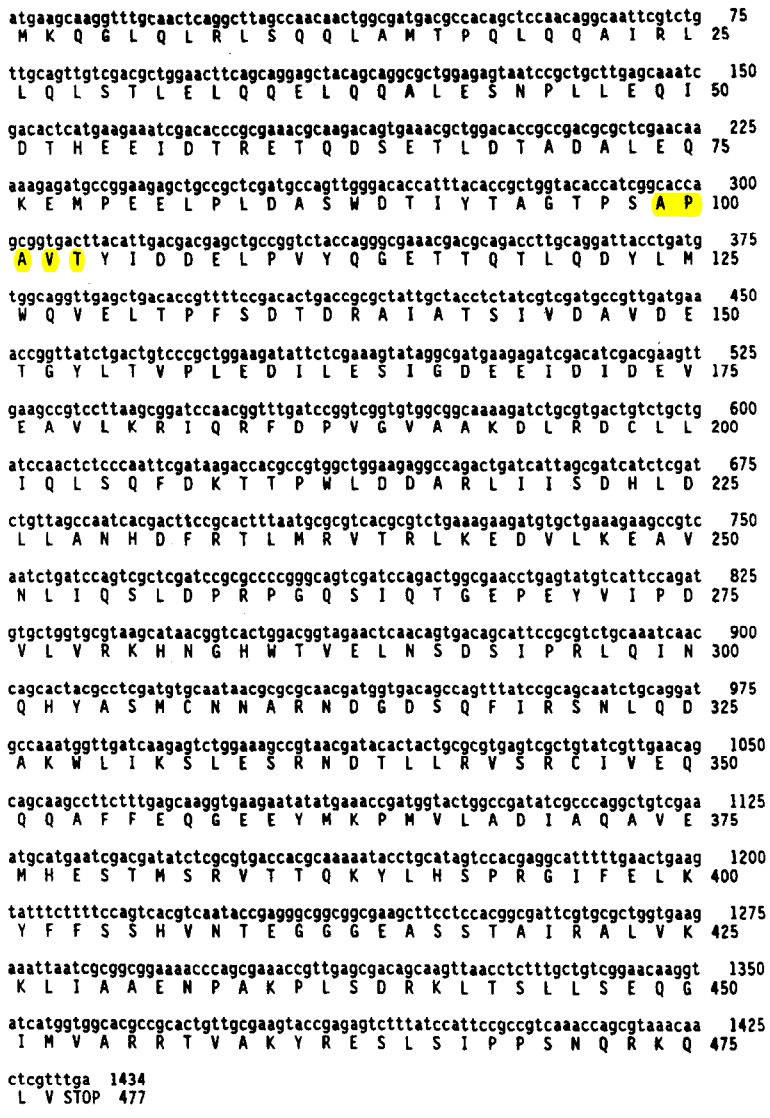

Figure 8. Sequence of the E. coli σ^{54} Gene

How Does Melting Occur?

The key activation event is DNA melting within a closed complex. Closed complex formation requires the C-terminus of the protein containing an HTH domain (see the model in Figure 7). Within this closed complex, the intramolecular leucine zipper forms an essential positioning domain near the DNA region that will be melted upon activation. This constrains the essential acid domain, which connects the two halves of the leucine zipper, to be near the region of DNA that will be melted. The melting only occurs in response to an unknown signal from the phosphorylated NR_I activator when it binds to its upstream sites. Since the acid domain is required for DNA melting but not promoter recognition, it must respond to this signal by helping to melt the DNA region over which it is positioned.

The data do not, however, indicate whether the acidic region melts the DNA directly, as implied in this model, or is simply part of the internal signal transduction pathway that leads to DNA melting. If the acid region acted indirectly by binding to NR_I one might expect a strong basic region on that protein; however, none is present. Direct melting in response to an NR_I-triggered conformational change would bring the negatively charged acid region nearer to the DNA region to be melted. There it could strip shielding cations and combine with the two negatively charged DNA strands to form an unstable pocket of very high negative charge density. One way to relieve this unfavorable electrostatic repulsion would be for the DNA strands to separate, with the template strand being transferred to its required receptor site within the catalytic domain on core polymerase. In this hypothetical model the acid domain acts as an electrostatic melter of the promoter.

Experimental Procedures

Strains and Materials

The plasmid pTH7 (Figure 1) was kindly provided by Dr. Boris Magasanik. It contains the E. coli σ^{54} gene cloned behind the *tac* promoter on a pBR322 derivative (Hunt and Magasanik, 1985). The background

strain, SDKTI (*galE*$^-$, Δ(lac,pro), fis::Tn5, *rpoN*::Tn10/ F′, I^q,*lacZU118, proAB*), was prepared by transducing RJ161 (F$^-$, *galE*, Δ(lac,pro); kindly provided by Dr. Reid Johnson) with P1 lysate from CSH26, *fis*::Tn5 (kindly provided by Catherine Ball). The resulting strain was then transduced with P1 lysate from YMC18 (*thi, endA, hsr,* ΔlacU169, rpoN::Tn10), a strain kindly provided by Dr. Boris Magasanik. The new strain was then mated with RJ184 (CSH26, *recA56*/F′ I^q, *lacZU118, proAB*; kindly provided by Dr. Reid Johnson), to obtain SDKTI carrying the I^q gene on an episome and a transposon inserted into the chromosomal σ^{54} gene. In addition to SDKTI, the strain SD16T (JM105, *rpoN*::Tn10) was created by transduction of JM105 with the lysate described above. Both of these strains contain an episome with the *lacI*^{q} gene and a Tn10 transposon that interrupts the σ^{54} gene and were therefore used interchangeably for the mutant screening and footprinting analyses.

All primers used for sequencing were synthesized at the UCLA Synthesizer Facility. Klenow fragment of DNA polymerase I was obtained from Bethesda Research Laboratories, Taq polymerase was obtained from New England Biolabs, and nuclease Bal31 was obtained from US Biochemicals. The Amersham in vitro mutagenesis kit, version 2.0 (code RPN.1523), was used for the acid deletion experiment.

Sequencing

The σ^{54} gene carried on the pTH7 plasmid and the pTH7 deletion plasmids were sequenced by double-stranded alkaline denaturation (Haltiner et al., 1985) using ^{32}P 5′ end–labeled primers, with modifications as described previously (Borowiec et al., 1987). The primers used for sequencing the deletion plasmids were the same used to sequence the gene. Both DNA strands were sequenced. The mutant plasmids were purified according to the method of Holmes and Quigley (1981) or that of Morelle (1989) for restriction analysis and sequencing. The sequence of the σ^{54} gene is shown in Figure 8.

Deletion Mutagenesis

Deletions were made at various positions within the *rpoN* gene either by use of the Amersham in vitro mutagenesis kit (see "Strains and Materials") or by cutting at the desired unique restriction site within pTH7 and chewing back both ends of the double-stranded DNA with nuclease Bal31 (Maniatis et al., 1988). In the latter procedure, 10 μg of pTH7 plasmid DNA was digested to completion with the desired restriction enzyme. The DNA was then extracted and precipitated. Nuclease Bal31 (1.5–5.0 U) was used to treat the digested DNA for 1.5–10 min. The DNA was again extracted, precipitated, and end-filled with Klenow fragment (Maniatis et al., 1988). Approximately 0.1 μg of linear plasmid DNA carrying the collection of deletions was blunt-end ligated and transformed into the strain SDKTI or SD16T. Colonies (48–123) were then screened for growth on W-Gln, W-Gln containing IPTG, and W-Arg plates supplemented with kanamycin (40 μg/ml), tetracycline (5 μg/ml), and ampicillin (100 μg/ml). In preparing these plates, W-salts media (a minimal medium lacking nitrogen) were prepared as described previously (Smith et al., 1971; Reitzer and Magasanik, 1985) and supplemented with the appropriate antibiotics and either glutamine (0.2%) or arginine (0.01 mg/ml). IPTG was then added to a final concentration of 0.5 mM when desired. The screening technique was the same regardless of the mutagenesis procedure used.

In Vivo Footprinting

Cells (SDKTI or SD16T) carrying the desired σ^{54} deletion plasmid were grown overnight in the activating media GN-Gln (Smith et al., 1971; Reitzer and Magasanik, 1985) in the presence of ampicillin (100 μg/ml) and diluted 1:100 the next morning into 25 ml of the same media. When the cells reached 0.35–0.5 OD_{660}, 25 μl of 0.5 M IPTG was added to induce expression of the *rpoN* gene from the pTH7 plasmid. At ~0.7–0.8 OD_{660}, the cells were split into two 10 ml samples. One of these was treated with 50 μl of 40 mg/ml rifampicin (dissolved in methanol) for 5 min, followed by 270 μl of 0.37 M $KMnO_4$ for 2 min. The other was treated with 2.83 μl of 10.6 M DMS for 5 min. The cells were then immediately spun down and the chromosomal DNA was isolated according to the procedure of Owen and Borman (1987), with the following modifications. Following the addition of proteinase K, the samples were incubated overnight at 37–52°C rather than for 30 min at 37°C. In addition, following the initial extraction with chloroform, the DNA was extracted a second time with chloroform, once with phenol, and finally, once again with chloroform:isoamyl alcohol. Following precipitation, the DMS samples were cleaved with piperidine as described previously (Borowiec et al., 1987) and purified through a 1 ml G 50–80 Sephadex spin column in water. In contrast, the $KMnO_4$-treated samples were resuspended in 300 μl of water, and 200 μl of this was passed through the same type of column for use in primer extension analysis.

Primer Extension Analysis

In vivo sample (35 μl) or ~0.1–0.3 μg of DNA modified in vitro, as described previously (Gralla, 1985; Borowiec et al., 1987), was analyzed by a modification of the PCR (Mullis and Faloona, 1987). Briefly, 49.5 μl of water, 10 μl of 10× Taq polymerase buffer (166 mM $[NH_4]_2SO_4$, 0.67 M Tris [pH 8.8], 67 mM $MgCl_2$, 100 mM β-mercaptoethanol, 1 mg/ml bovine serum albumin), 4 μl of 5 mM dNTP mix, 1 μl of ^{32}P 5′end–labeled *glnD* primer (Keener and Kustu, 1988) (500 × 10^5 cpm/μl), 35 μl of DNA solution and 0.5 μl (2.5 U/μl) of Taq polymerase were mixed together and covered with 100 μl of mineral oil. These samples were then denatured at 94°C for 1.5 min, hybridized 2′ at 57°C, and extended 3′ at 72°C. This cycle was repeated an additional 12–18 times except that the length of the denaturation was reduced to 1 min and the final cycle was extended for 10 min rather than 3 min. The samples were then removed from the mineral oil, extracted with chloroform:isoamyl alcohol, and precipitated with one-third volume of 4 M NH_4OAc/ 20 mM EDTA. The samples were then electrophoresed on a 6% denaturing polyacrylamide gel and an autoradiograph of the gel was obtained overnight.

Acknowledgments

We are very grateful to Dr. Boris Magasanik for his kind gifts of strains and advice, to Dr. Reid Johnson and Catherine Ball for strains, and to Caine Wong for his advice and help in proofreading. This research was supported by US Public Health Service grant GM35754. S. S. D. was supported by USPHS postdoctoral fellowship CA09056.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *"advertisement"* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received April 13, 1990; revised June 28, 1990.

References

Axelrod, J. D., and Majors, J. (1989). An improved method for photofootprinting yeast genes *in vivo* using Taq polymerase. Nucl. Acids Res. *17*, 171–183.

Birkmann, A., and Böck, A. (1989). Characterization of a *cis* regulatory DNA element necessary for formate induction of the formate dehydrogenase gene (*fdh F*) of *E. coli*. Mol. Microbiol. *3*, 187–195.

Birkmann, A., Zinoni, F., Sawers, G., and Böck, A. (1987). Factors affecting transcriptional regulation of the formate-hydrogen-lyase pathway of *E. coli*. Arch. Microbiol. *148*, 44–51.

Borowiec, J. A., Zhang, L., Sasse-Dwight, S., and Gralla, J. D. (1987). DNA supercoiling promotes formation of a bent repression loop in *lac* DNA. J. Mol. Biol. *196*, 101–111.

Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A., and Sauer, R. T. (1990). Deciphering the message in protein sequences: tolerance to amino acid substitutions. Science *247*, 1306–1310.

Brendel, V., and Karlin, S. (1989). Association of charge clusters with functional domains of cellular transcription factors. Proc. Natl. Acad. Sci. USA *86*, 5698–5702.

Brennan, R. T., and Matthews, B. W. (1989). The helix-turn-helix DNA binding motif. J. Biol. Chem. *264*, 1903–1906.

Buratowski, S., Hahn, S., Guarente, L., and Sharp, P. A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. Cell *56*, 549–561.

Busch, S. J., and Sassone-Corsi, P. (1990). Dimers, leucine zippers and DNA binding domains. Trends Genet. *6*, 36–40.

Courey, A. J., and Tjian, R. (1988). Analysis of Sp1 in vivo reveals multi-

ple transcriptional domains, including a novel glutamine-rich activation motif. Cell *55*, 887–898.

Courey, A. J., Holtzman, D. A., Jackson, S. P., and Tjian, R. (1989). Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. Cell *59*, 827–836.

Gardella, T., Moyle, H., and Susskind, M. (1989). A mutant *Escherichia coli* σ^{70} subunit of RNA polymerase with altered promoter specificity. J. Mol. Biol. *206*, 579–590.

Giniger, E., and Ptashne, M. (1987). Transcription in yeast activated by a putative amphipathic alpha helix linked to a DNA binding unit. Nature *330*, 670–672.

Gralla, J. D. (1985). Rapid "footprinting" of supercoiled DNA. Proc. Natl. Acad. Sci. USA *82*, 3078–3081.

Gralla, J. D. (1989a). Bacterial gene regulation from distant DNA sites. Cell *57*, 193–195.

Gralla, J. D. (1989b). Promoter recognition and mRNA initiation by *E. coli* $E\sigma^{70}$. Meth. Enzymol. *185*, 37–54.

Haltiner, M., Kempe, T., and Tjian, R. (1985). A novel strategy for constructing clustered point mutations. Nucl. Acids Res. *13*, 1015–1025.

Holmes, D. S., and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. *114*, 193–197.

Hope, I. A., Mahadevan, S., and Struhl, K. (1988). Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. Nature *333*, 635–640.

Hunt, T. P., and Magasanik, B. (1985). Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG* and *glnL*. Proc. Natl. Acad. Sci. USA *82*, 8453–8457.

Keener, J., and Kustu, S. (1988). Protein kinase and phosphoprotein phosphotase activities of nitrogen regulatory proteins NtrB and NtrC of enteric bacteria: roles of the conserved amino-terminal domain of NtrC. Proc. Natl. Acad. Sci. USA *85*, 4976–4980.

Landshulz, W. H., Johnson, P. F., and McKnight, S. L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science *240*, 1759–1764.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1988). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Merrick, M. J., and Gibbins, J. R. (1985). The nucleotide sequence of the nitrogen-regulation gene *ntrA* of *Klebsiella pneumoniae* and comparison with conserved features in bacterial RNA polymerase sigma factors. Nucl. Acids Res. *13*, 7607–7620.

Merrick, M., Gibbins, J., and Toukdarian, A. (1987). The nucleotide sequence of the sigma factor gene *ntrA (rpoN)* of *Azotobacter vinelandii*: analysis of conserved sequences in NtrA proteins. Mol. Gen. Genet. *210*, 323–330.

Morelle, G. (1989). A plasmid extraction procedure on a miniprep scale. Focus *11*, 7–8.

Mullis, K. B., and Faloona, F. A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Meth. Enzymol. *155*, 335–350.

Nakajima, N., Horikoshi, M., and Roeder, R. (1988). Factors involved in specific transcription by mammalian RNA polymerase. II. Purification, genetic specificity and TATA box–promoter interactions of TFIID. Mol. Cell. Biol. *8*, 4028–4040.

Ninfa, A. J., and Magasanik, B. (1986). Covalent modification of the *glnG* product NR_I, by the *glnL* product, NR_{II}, regulates transcription of the *glnALG* operon in *E. coli*. Proc. Natl. Acad. Sci. USA *83*, 5909–5913.

Ninfa, A. J., Reitzer, L. J., and Magasanik, B. (1987). Initiation of transcription at the bacterial *glnAp2* promoter by purified E. coli components is facilitated by enhancers. Cell *50*, 1039–1046.

O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989a). Evidence that the leucine zipper is a coiled coil. Science *243*, 538–542.

O'Shea, E. K., Rutkowski, R., Stafford, W. F., III, and Kim, P. S. (1989b). Preferential heterodimer formation by isolated leucine zippers from *fos* and *jun*. Science *245*, 646–648.

Owen, R. J., and Borman, P. (1987). A rapid biochemical method for purifying high molecular weight bacterial chromosomal DNA for restriction analysis. Nucl. Acids Res. *15*, 3631.

Popham, D. L., Szeto, D., Keener, J., and Kustu, S. (1989). Function of a bacterial activator protein that binds to transcriptional enhancers. Science *243*, 629–635.

Reitzer, L. J., and Magasanik, B. (1985). Expression of *glnA* in *Escherichia coli* is regulated at tandem promoters. Proc. Natl. Acad. Sci. USA *82*, 1979–1983.

Reitzer, L. J., and Magasanik, B. (1986). Transcription of *glnA* in E. coli is stimulated by activator bound to sites far from the promoter. Cell *45*, 785–792.

Ronson, C. R., Nixon, B. T., Albright, L., and Ausubel, F. M. (1987). *Rhizobium meliloti ntrA (rpoN)* gene is required for diverse metabolic functions. J. Bacteriol. *169*, 2424–2431.

Sasse-Dwight, S., and Gralla, J. D. (1988). Probing the *glnALG* upstream activation mechanism *in vivo*. Proc. Natl. Acad. Sci. USA *85*, 8934–8938.

Sasse-Dwight, S., and Gralla, J. D. (1989). $KMnO_4$ as a probe for *lac* DNA melting and mechanism *in vivo*. J. Biol. Chem. *264*, 8074–8081.

Siegele, D. A., Hu, J. C., Walter, W. A., and Gross, C. A. (1989). Altered promoter recognition by mutant forms of the σ^{70} subunit of *Escherichia coli* RNA polymerase. J. Mol. Biol. *206*, 591–603.

Smith, G. R., Halpern, Y. S., and Magasanik, B. (1971). Genetic and metabolic control of enzymes responsible for histidine degradation in *S. typhimurium*. J. Biol. Chem. *246*, 3320–3329.

Sorger, P. K., and Nelson, H. C. M. (1989). Trimerization of a yeast transcriptional activator via a coiled-coil motif. Cell *59*, 807–813.

Vinson, C. R., Sigler, P. B., and McKnight, S. L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science *246*, 911–916.

Yudkin, M. D. (1987). Structure and function of a *Bacillus subtilis* sporulation-specific sigma factor: molecular nature of mutations in SpoIIAC. J. Gen. Microbiol. *133*, 475–481.

GenBank Accession Number

The accession number for the sequence reported in this paper is M36622.